



AF  
2655

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): Y. Tahara et al.  
Docket No.: JP919990203  
Serial No.: 09/656,963  
Filing Date: September 7, 2000  
Group: 2655  
Examiner: Michael N. Opsasnick

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Signature: Lisa L. Vulpis Date: May 25, 2004

Title: Methods and Apparatus for Voice Information Registration  
and Recognized Sentence Specification in Accordance  
With Speech Recognition

TRANSMITTAL OF APPEAL BRIEF

RECEIVED

JUN 03 2004

Technology Center 2600

Mail Stop Appeal Brief - Patents  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Submitted herewith are the following documents relating to the above-identified patent application:

- (1) Response to Office Action; and
- (2) Supplemental Appeal Brief in triplicate (original and two copies).

There is no additional fee due in conjunction with the response. In the event of non-payment or improper payment of a required fee, the Commissioner is authorized to charge or to credit **International Business Machines Corporation Deposit Account No. 50-0510** as required to correct the error.

Respectfully submitted,

Robert W. Griffith

Date: May 25, 2004

Robert W. Griffith  
Reg. No. 48,956  
Attorney for Applicant(s)  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-4547



Attorney Docket No. JP919990203

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): Y. Tahara et al.  
Docket No.: JP919990203  
Serial No.: 09/656,963  
Filing Date: September 7, 2000  
Group: 2655  
Examiner: Michael N. Opsasnick

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Signature: *Linda L. Tulpis*

Date: May 25, 2004

Title: Methods and Apparatus for Voice Information Registration  
and Recognized Sentence Specification in Accordance  
With Speech Recognition

RECEIVED

JUN 03 2004

SUPPLEMENTAL APPEAL BRIEF

Technology Center 2600

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313

Sir:

This Supplemental Appeal Brief is submitted in response to the Office Action dated February 25, 2004 in the above-referenced application, in which the Examiner reopened prosecution in response to the Appeal Brief filed October 6, 2003.

Appellants have submitted concurrently herewith a response to the Office Action, requesting reinstatement of the appeal pursuant to 37 C.F.R. §1.193(b)(2).

REAL PARTY IN INTEREST

The present application is assigned to International Business Machines Corp., as evidenced by an assignment recorded December 5, 2000 in the U.S. Patent and Trademark Office at Reel 11315, Frame 0734. The assignee, International Business Machines Corp., is the real party in interest.

## RELATED APPEALS AND INTERFERENCES

There are no known related appeals and interferences.

## STATUS OF CLAIMS

Claims 1-15 are pending in the present application. Claims 5, 10 and 15 are allowable, and claims 1-4, 6-9 and 11-14 stand rejected under 35 U.S.C. §103(a). Claims 1-4, 6-9 and 11-14 are appealed.

## STATUS OF AMENDMENTS

There have been no amendments filed subsequent to the rejection. However, a Response to Final Office Action was filed July 31, 2003, and an Appeal Brief was filed on October 6, 2003.

## SUMMARY OF INVENTION

The present invention relates to a word registration method for a speech recognition system and, more particularly, to a method whereby voice is used to specify information displayed on a screen (Specification, page 1, lines 5-8).

By way of example, as recited in claim 1, a voice information registration method, employed by a speech recognition apparatus, comprises the following steps. A sentence group is obtained, which includes a first to an  $N$ -th sentence, wherein  $N$  is a number equal to or greater than two. A sounds-like spelling is obtained for a word that is included in an  $i$ -th sentence, but is not entered in a speech recognition dictionary, wherein  $i$  is a number equal to or less than  $N$ . A base form is obtained based on the sounds-like spelling of the word. Finally, the base form is registered in a speech recognition dictionary in correlation with the word.

In an illustrative embodiment, a group of sentences to be recognized is obtained from an application. Using parsing logic, each target sentence to be recognized is divided into words, or speech recognition units. Thereafter, the words in each target sentence are examined to determine whether there are unknown words among them that are not registered in the speech recognition dictionary, but for which the sounds-like spelling is available. If an unknown word is found, a base form, from which the pronunciation is inferred from the sounds-like spelling, is prepared and

registered in the speech recognition dictionary. This base form is employed when the voice of a user is recognized who has orally designated one of the sentences (Specification, page 2, lines 4-12)

According to one aspect of the present invention, a voice information registration method is provided, which is employed by a speech recognition apparatus, and with which a voice input device is used (Specification, page 2, lines 13-15). According to another aspect of the present invention, a sentence specification method is provided that is employed by a speech recognition apparatus, and with which a voice input device is used. This sentence specification method has both a registration step and a recognition step (Specification, page 2, lines 24-26).

An unknown word, detected in an extracted sentence, is recognized as a word but is not registered in the speech recognition dictionary. Thus the base form of the unknown word is unknown to the system (Specification, page 17, lines 5-7). A flow diagram showing the unknown word detection processing performed according to an embodiment of the present invention is shown in FIG. 5. A flow diagram showing the processing performed according to an embodiment of the present invention to obtain a base form corresponding to an unknown word and to register the base form in a speech recognition dictionary is shown in FIG. 6. As described in the figures, the present invention allows for the recognition of a sentence that includes words that are not registered in a speech recognition dictionary through the registration process of the present invention (Specification, page 26, lines 10-12).

#### ISSUE PRESENTED FOR REVIEW

Whether claims 1-4, 6-9 and 11-14 are properly rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 5,208,897 to Hutchins (hereinafter "Hutchins"), in view of U.S. Patent Application Publication No. 2003/0229497 to Wilson et al. (hereinafter "Wilson"), and IBM Technical Disclosure Bulletin, Vol. 35, Issue 1a, p. 59 (hereinafter "TDB").

#### GROUPING OF CLAIMS

Claims 1-4, 6-9 and 11-14 stand or fall together.

### ARGUMENT

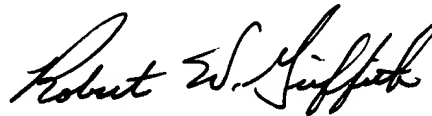
Appellants incorporate by reference herein the disclosure of all previous responses filed in the present application, namely, responses dated February 19, 2003 and July 31, 2003, and an Appeal Brief filed October 6, 2003.

With regard to the issue of whether claims 1-4, 6-9 and 11-14 are unpatentable under 35 U.S.C. §103(a) over Hutchins in view of Wilson and IBM TDB, Appellants respectfully assert that claims 1-4, 6-9 and 11-14 are not properly rejected under 35 U.S.C. §103(a).

Appellants submit that Wilson, which is used in combination with Hutchins and IBM TDB, is not prior art. Wilson has a filing date of December 11, 2003 and is a continuation-in-part of an application filed on April 21, 2000. No evidence has been provided indicating that the portions of Wilson relied upon by the Examiner are entitled to receive the benefit of the earlier filing date. Further, although the present application has a U.S. filing date of a September 7, 2000, it has claimed foreign priority benefits under 35 U.S.C. §119 based on Japanese Application No. 11-257587 filed on September 10, 1999. This foreign priority was acknowledged by the Examiner in an Office Action dated November 19, 2002. Therefore, the effective filing date of the present invention is September 10, 1999, before the earliest possible effective filing date of Wilson (April 20, 2000), which precludes Wilson as a prior art reference.

For at least the reason presented above, Appellants respectfully request withdrawal of the §103(a) rejection of claims 1-4, 6-9 and 11-14. As such, the application is asserted to be in condition for allowance, and favorable action is respectfully solicited.

Respectfully submitted,



Date: May 25, 2004

Robert W. Griffith  
Attorney for Applicant(s)  
Reg. No. 48,956  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-4547

## APPENDIX

1. A voice information registration method, employed by a speech recognition apparatus, comprising:

(a) obtaining a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two;

(b) obtaining a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N;

(c) obtaining a base form based on said sounds-like spelling of said word; and

(d) registering said base form in a speech recognition dictionary in correlation with said word.

2. A sentence specification method, employed by a speech recognition apparatus, comprising:

a registration step including:

(a1) obtaining a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two,

(a2) obtaining a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N,

(a3) obtaining a base form based on said sounds-like spelling of said word, and

(a4) registering said base form in a speech recognition dictionary in correlation with said word, and

a recognition step including:

(b1) obtaining voice information that is input as a user reads and vocally reproduces a display corresponding to said i-th sentence,

(b2) employing said base form to recognize said voice information and to select a speech recognition sentence, and

(b3) comparing said i-th sentence with said selected speech recognition sentence.

3. The sentence specification method according to claim 2, wherein said group of sentences is obtained from an application, said method further comprising a step of generating a control message corresponding to said i-th sentence and transmitting said control message to said application.

4. The sentence specification method according to claim 2, wherein a sounds-like spelling score is stored in correlation with the sounds-like spelling of said word, wherein a pronunciation score is stored in correlation with said base form, and wherein, when a function value that is obtained by using said sounds-like spelling score and said pronunciation score exceeds a threshold value, said base form is registered in a speech recognition dictionary.

6. A speech recognition apparatus, comprising:

(a) a sentence specification unit for obtaining a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two;

(b) an unknown word detector for obtaining a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N;

(c) a base form generator for obtaining a base form based on said sounds-like spelling of said word; and

(d) a speech recognition dictionary to which said base form is stored in correlation with said word.

7. A speech recognition apparatus, comprising:

(a) a sentence specification unit for obtaining a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two;

(b) an unknown word detector for obtaining a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N;

(c) a base form generator for obtaining a base form based on said sounds-like spelling of said word;

(d) a speech recognition dictionary in which said base form is stored in correlation with said word;

(e) a voice input unit for obtaining voice information that is input as a user reads and vocally reproduces a display corresponding to said i-th sentence; and

(f) a speech recognition engine for employing said base form to recognize said voice information and to select a speech recognition sentence;

wherein said sentence specification unit compares said i-th sentence with said selected speech recognition sentence.

8. The speech recognition apparatus according to claim 7, wherein said sentence specification unit obtains said group of sentences from an application, generates a control message corresponding to said i-th sentence, and transmits said control message to said application.

9. The speech recognition apparatus according to claim 7, wherein a sounds-like spelling score is stored in correlation with the sounds-like spelling of said word, wherein a pronunciation score is stored in correlation with said base form, and wherein, when a function value that is obtained by using said sounds-like spelling score and said pronunciation score exceeds a threshold value, said base form is registered in a speech recognition dictionary.

11. A storage medium in which a program for specifying a sentence is stored to be executed by a speech recognition apparatus, said program comprising:

(a) program code for instructing said speech recognition apparatus to obtain a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two;

(b) program code for instructing said speech recognition apparatus to obtain a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N;



(c) program code for instructing said speech recognition apparatus to obtain a base form based on said sounds-like spelling of said word; and

(d) program code for instructing said speech recognition apparatus to register said base form in a speech recognition dictionary in correlation with said word.

12. A storage medium in which a program for specifying a sentence is stored to be executed by a speech recognition apparatus, said program comprising:

(a) program code for instructing said speech recognition apparatus to obtain a sentence group, which includes a first to an N-th sentence, wherein N is a number equal to or greater than two;

(b) program code for instructing said speech recognition apparatus to obtain a sounds-like spelling for a word that is included in an i-th sentence, but is not entered in a speech recognition dictionary, wherein i is a number equal to or less than N;

(c) program code for instructing said speech recognition apparatus to obtain a base form based on said sounds-like spelling of said word;

(d) program code for instructing said speech recognition apparatus to register said base form in a speech recognition dictionary in correlation with said word;

(e) program code for instructing said speech recognition apparatus to obtain voice information that is input as a user reads and vocally reproduces a display corresponding to said i-th sentence;

(f) program code for instructing said speech recognition apparatus to employ said base form to recognize said voice information and to select a speech recognition sentence; and

(g) program code for instructing said speech recognition apparatus to compare said i-th sentence with said selected speech recognition sentence.

13. The storage medium according to claim 12, wherein said group of sentences is obtained from an application, and wherein program code is stored to instruct said speech recognition apparatus to generate a control message corresponding to said i-th sentence and to transmit said control message to said application.

14. The storage medium according to claim 12, wherein a sounds-like spelling score is stored in correlation with the sounds-like spelling of said word, wherein a pronunciation score is stored in correlation with said base form, and wherein, when a function value that is obtained by using said sounds-like spelling score and said pronunciation score exceeds a threshold value, said based form is registered in a speech recognition dictionary.